The eXperanto Experiment: Effects of Explicit Instruction on Second Language Acquisition

de Graaff, H.C.J.

published in
Studies in Second Language Acquisition
1997

document version
Publisher's PDF, also known as Version of record

SSLA, **19**, 249–276. Printed in the United States of America.

THE *EXPERANTO* EXPERIMENT

Effects of Explicit Instruction on Second Language Acquisition

Rick de Graaff
Vrije Universiteit Amsterdam

Theories on the role of consciousness and the enhancement of noticing (Schmidt, 1990, 1994; Sharwood Smith, 1993) predict a facilitative effect of explicit knowledge, as built up by explicit instruction, on the acquisition of implicit second language (L2) knowledge. This study investigates the interaction between the presence or absence of explicit instruction and the variables complexity and morphology/syntax in the acquisition of four L2 structures. Two groups of 27 university students, differing in the exposure to explicit instruction, followed a computer-controlled self-study course in an artificial language. Results from computer-controlled posttests confirm the general hypothesis that explicit instruction facilitates the acquisition of L2 grammar. However, no evidence could be reported for the hypotheses predicting a differential effect of explicit instruction depending on the variables complexity and morphology/syntax.

One of the most persistent questions in second language (L2) acquisition research and pedagogy concerns the impact that explicit instruction of grammar may have on L2 acquisition. Implicit language knowledge is generally considered to form the basis for L2 performance. However, the way in which this implicit knowledge is built up most effectively is still an issue of considerable disagreement. Essentially, there

This article is based on a dissertation project titled "The Influence of Grammar Instruction on Foreign Language Acquisition," which is supported by the Foundation for Language, Speech and Knowledge (under grant 300-73-028), funded by the Netherlands Organization for Scientific Research (NWO). I thank Jan Hulstijn, Peter Jordens, Huub van den Bergh, Robert DeKeyser, Peter Robinson, Mitsuhiko Ota, and an anonymous *SSLA* reviewer for their comments on an earlier version of this article.

Address correspondence to Rick de Graaff, Vrije Universiteit, Department of Applied Linguistics, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, e-mail: graaff@jet.let.vu.nl.

are two points of view in this debate. One is that implicit knowledge cannot be acquired as the result of explicit learning—for example, by means of explicit instruction—nor can it indirectly be influenced by explicit knowledge of the L2. This point of view, which has become known as the noninterface position, has most strongly been advocated by Krashen (1981, 1982, 1985). Among researchers working within Universal Grammar theory, a similar position has been taken by, for example, Schwartz (1993). The other point of view is that implicit L2 knowledge can be acquired through or facilitated by explicit knowledge. In its strongest form, this interface position considers explicit, declarative knowledge to be convertible into implicit, procedural knowledge through practice. This strong version of the interface position is based on literature in cognitive psychology, in particular, theories of controlled and automatic processing (e.g., Anderson, 1982; Shiffrin & Schneider, 1977). Although this position has not been very common in empirical L2 acquisition research, several authors have invoked a strong interface model in nonempirical studies (e.g., DeKeyser, in press; McLaughlin, 1990; McLaughlin, Rossman, & McLeod, 1983; O'Malley & Chamot, 1990). Furthermore, in L2 pedagogy it is common practice in many cases to teach L2 structures through gradual automatization of explicit knowledge. A weaker form of the interface position holds that explicit knowledge as built up by explicit instruction or any other input-enhancing technique (Sharwood Smith, 1993) helps learners to notice crucial properties of the input (Schmidt, 1994; Tomlin & Villa, 1994), thus facilitating the process of intake (R. Ellis, 1993) and the subsequent acquisition of implicit knowledge. This focus on form (Long, 1991; Long & Robinson, in press) has been claimed to be especially effective when it strengthens form–meaning relationships (Terrell, 1991; VanPatten, 1994).

A considerable number of experimental studies using a wide range of different research designs have been carried out in recent years in order to gain insight into the assumed facilitative effect of explicit instruction. Experimental settings have been used by, among others, Alanen (1995), Carroll and Swain (1993), DeKeyser (1995), Doughty (1991), Robinson (1996, this issue), and Yang and Givón (1993); see Hulstijn (this issue) for a more complete overview. Other researchers have studied the effect of explicit instruction in classroom settings (e.g., van Baalen, 1983; Cadierno, 1995; DeKeyser, 1993; Elliot, 1995; Fotos, 1993; Harley, 1989; Herron & Tomasello, 1988; Lightbown & Spada, 1990; Pica, 1983; Scott, 1989; VanPatten & Cadierno, 1993; VanPatten & Sanz, 1995; Yan-Ping, 1991). Most studies, in both experimental and classroom settings, used one or more experimental groups and a control group. Learners in experimental groups received either only explicit instruction (DeKeyser, 1995; Fotos, 1993; Robinson, 1996; Scott, 1989; Yan-Ping, 1991) or explicit instruction plus practice (van Baalen, 1983; Doughty, 1991; Pica, 1983), only explicit feedback (Carroll & Swain, 1993; DeKeyser, 1993), explicit feedback plus practice (Herron & Tomasello, 1988), or explicit "input processing" (Cadierno, 1995; VanPatten & Cadierno, 1993; VanPatten & Sanz, 1995). In most studies, treatment was short (one or two sessions) and testing was carried out immediately after treatment. Some studies were conducted over a longer treatment period and comprised additional delayed posttests. A clear positive effect of some kind of explicit instruction was reported by Cadierno (1995), Carroll and Swain (1993), Doughty (1991), DeKeyser (1995),

Herron and Tomasello (1988), Robinson (1996), Scott (1989), VanPatten and Cadierno (1993), and VanPatten and Sanz (1995). However, because the studies listed here comprise a wide range of research designs, experimental conditions, test types, languages, age groups, and learning contexts, it is difficult to directly compare their findings.

Although some evidence for the facilitative effect of explicit instruction on L2 acquisition has been found, little is known yet concerning the question of under which specific learning circumstances and for exactly which aspects of grammar explicit knowledge can be most facilitative for L2 acquisition. In Hulstijn and de Graaff (1994), we argued that the following variables could possibly influence the effect of explicit instruction: the target structure's linguistic domain (in the core or in the periphery of Universal Grammar), its complexity (see below), its degree of semantic redundancy, its reliability (the ratio of regular to irregular cases), its scope (number of phenomena covered), the frequency with which it is manifested in the input, the competition between rule-based learning and item-based learning (see below), task modality (reception vs. production), type of instruction or input enhancement, and individual learner characteristics. Some of these variables are featured in the present study and will be briefly discussed in the next sections.

Complexity and Difficulty of Grammar Structure

As has already been mentioned, there is some evidence that explicit instruction facilitates the acquisition of implicit L2 knowledge. The weak interface position hypothesizes this is due to the enhancement of noticing (Schmidt, 1990). This implies that explicit instruction will be more effective in the case of target structures that are not perceptually salient (cf. Robinson, 1996; see Slobin's [1985] operating principles for a framework for the relationship between perceptual and cognitive salience and learning). When target structures can be noticed and processed from the input spontaneously, explicit instruction is less necessary and, in consequence, its effect will be smaller. The likelihood of spontaneous noticing and processing may depend on, among other things, the complexity of the target structure.

A few studies to date have tried to empirically or theoretically define the concept of complexity. Van Baalen (1983) ordered target structures in terms of complexity by means of teacher and textbook judgments. Robinson (1996), too, used judgments of experienced L2 teachers to identify the complexity of the pedagogic rules under investigation. Krashen (1982) distinguished easy and hard rules, claiming that explicit instruction can only be helpful in the case of easy rules, because only those are explicitly teachable and learnable. Hulstijn and de Graaff (1994) defined complexity as the number of grammatical concepts that have to be taken into account in order to arrive at a correct form in language production (p. 103). When more criteria have to be applied, explicit instruction is hypothesized to be more effective, as spontaneous noticing and processing would then be less likely. Hulstijn and de Graaff's definition of complexity obviously does not predict ease of acquisition as such, because this also depends on the amount of processing capacity that the criteria in question require, as well as on the transparency or opacity of the form–

meaning relationship established by the structure (R. Ellis, 1990). Furthermore, R. Ellis (1994) speculated that explicit instruction directed at too complex a structure is likely to lead only to improved accuracy in planned language use (p. 623). This distinction might be what guided Krashen (1981) in his claim concerning the effect of explicit instruction on easy as opposed to hard rules, the latter referring to structures that learners are not "ready" to acquire yet.

Rule-Based Versus Item-Based Learning

Implicit L2 knowledge can be built up by different types of learning processes. Using evidence from cognitive psychology, Carr and Curran (1994) claim that both abstract rule-based and exemplar-based representations of grammar can contribute to the acquisition of L2 knowledge (see Robinson, this issue, for a detailed discussion on the relationship between rule-based and item-based learning). DeKeyser (1995), in that respect, found no significant differences between abstract rule-based learning and exemplar-based item learning for categorical morphological rules applied to items that had been part of the learning set. For prototypical rules, some evidence was found that item-based learning was more effective than rule-based learning. For categorical rules, however, when applied to new forms, rule-based learning was more effective than item-based learning.

In syntax, as opposed to morphology, it can be hypothesized that exemplar-based item learning is less likely to occur. Whereas learners are likely to store individual, inflected word forms, syntactic structure depends on a larger context and for that reason has to be processed and acquired beyond item level. This would suggest that, apart from memorization of fixed chunks, acquisition of syntactic structure depends more heavily on abstract rule-based learning than the acquisition of morphological structure (Hulstijn & de Graaff, 1994).

Monitoring

Spontaneous, fluent language performance is unconscious in the sense that it is generated by implicit linguistic knowledge, without conscious retrieval of explicit knowledge (Schmidt, 1994). Under certain circumstances, however, explicit knowledge can be accessible during processing and can play an additional role in the production of language. The monitor theory (Krashen, 1985) claims that explicit knowledge is only accessible when enough time is available for conscious monitoring—for example, in written tasks. When processing takes place for spontaneous language use or under time pressure, explicit knowledge is not or much less accessible for processing. In empirical terms, this implies that tasks exerting high time pressure are more likely to elicit performance stemming exclusively from implicit knowledge than tasks allowing ample time.

Learner Characteristics

A wide range of variables has been distinguished in L2 acquisition research in order to explain individual differences among L2 learners (see Skehan, 1989, for an overview). Of these variables, language aptitude has been found to be one of the

best predictors of success in L2 learning (Carroll, 1981; R. Ellis, 1994; Gardner & MacIntyre, 1992). In L2 acquisition studies (e.g., DeKeyser, 1995; Robinson, 1995), language aptitude is usually measured by means of the Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1959), which aims at tapping both underlying language learning capacity and the capacity to handle decontextualized language (Skehan, 1989). Skehan suggests that the different components of language aptitude are relevant for both instructed and naturalistic language learning, as in both environments, consciously or unconsciously, processing of form and meaning is involved.

Although Carroll (1965) hypothesized that high-quality language instruction may nullify aptitude differences, there is no clear evidence that learners with low aptitude profit more from instruction than learners with high aptitude. Krashen (1981) even claimed that language aptitude can only predict proficiency differences in the case of (explicit) "learning," but not in the case of (implicit) "acquisition," which suggests that instruction would enlarge proficiency differences. Zobl (1992) argued that meaning-focused treatment without focus on form would result in smaller individual differences, indicating a weaker relationship between aptitude and meaning-focused learning than between aptitude and form-focused learning. Robinson (1995), however, found that aptitude is related to learning and awareness not only under instructed and rule-search conditions but also under implicit conditions. Only for learners under incidental learning conditions was no relationship found. These findings suggest that aptitude affects language learning under any circumstances for which students focus on form, which might be the case both with and without explicit instruction.

THE STUDY

Hypotheses

The present study investigated the interaction between instructional condition (presence or absence of explicit grammar instruction) and four other variables (two language structure variables, a processing variable, and a learner variable) in the acquisition of four L2 grammar structures. A computer-controlled learning setting and an artificial language were used in order to optimally control the input and exposure. The following hypotheses were tested.

Hypothesis 1. Participants receiving explicit instruction will perform better on tests measuring proficiency in the target structures than those not receiving explicit instruction. Language aptitude will affect test performance irrespective of the provision of explicit instruction.

This hypothesis is based on a weak version of the interface position that claims that explicit L2 knowledge, although not directly convertible into implicit L2 knowledge or L2 performance, facilitates language processing and thus facilitates the acquisition of implicit L2 knowledge. It is expected that language aptitude will affect learning to the same extent in both explicit and implicit instructional conditions, as they both require some kind of focus on form.

Hypothesis 2. Explicit instruction will be more effective in the case of the acquisition of complex structures than in the case of the acquisition of simple structures.

This hypothesis is motivated by the fact that simple formal phenomena may be clear enough in the input to be noticed and processed by L2 learners spontaneously, without explicit instruction. In the case of complex structures, however, explicit instruction may save learners considerable time and effort in discovering and processing their intricacies (Hulstijn & de Graaff, 1994). Complexity here is defined as the number of different grammatical concepts that have to be taken into account, explicitly or implicitly, in order to process a grammatical structure for language acquisition.

Hypothesis 3. Explicit instruction will be more effective in the case of the acquisition of syntactic structures than in the case of morphological structures.

This hypothesis addresses the fact that regular morphological structures can be acquired by explicit or implicit rule application as well as by analogy application (as a result of item-based learning), whereas for syntactic structures acquisition depends more exclusively on (explicit or implicit) rule application. Thus, in the case of morphological forms, two competing learning modes can be conceived of; in the case of syntactic structures, however, item-based learning is less available as an alternative to rule-based learning. This implies that explicit instruction in the case of syntactic structures will be relatively more influential than in the case of morphological structures (Hulstijn & de Graaff, 1994).

Hypothesis 4. Explicit instruction will be more effective when monitoring is possible during processing than when monitoring is inhibited by time pressure. However, under time pressure, a significant effect for instructional condition is expected as well.

This hypothesis is motivated by the applicability of explicit knowledge during task performance. Without time pressure, it is possible to rely on explicit knowledge for carrying out a linguistic task. Participants in the explicit condition are expected to build up a more structured and more complete explicit knowledge of the target structures. This will enable them to perform better on monitored tasks than participants in the implicit condition. Under time pressure, on the other hand, explicit knowledge is less accessible, and participants will rely more on implicit knowledge, which is why the weak interface position can be most directly tested here. As explicit instruction is expected to be facilitative for the acquisition of implicit knowledge, participants in the explicit condition are hypothesized to perform better also when explicit knowledge is less accessible. However, the difference between instructional conditions is expected to be greater without time pressure, as additional access to explicit knowledge then is more likely to occur.

Of the four hypotheses that were tested in this study, Hypothesis 1 is the most general one. The specific aim of the study is to provide evidence for the intervening role of other variables on the effect of explicit instruction. Hypotheses 2–4 aim to demonstrate the *differential* nature of Hypothesis 1. In all hypotheses, effectiveness of explicit instruction is measured as the proficiency *difference* between experimental groups.

Artificial Language and Laboratory Setting

The use of an artificial language in L2 acquisition research is basically motivated by reasons of experimental control. First, it ensures that participants cannot have contact with the target language nor receive any explicit grammar explanation except during the experimental sessions. Second, it provides the opportunity to filter out irregular or complex grammatical features or lexical items that are not the focus of the study, as these would otherwise undesirably complicate the learning task. At the same time, modifications can be made in order to make target structures exactly fit the linguistic requirements dictated by the research questions.

Artificial languages within this framework should be clearly distinguished from the artificial grammars as used in, for example, Reber's (1989) experiments. Such finite-state grammars consist of meaningless letter patterns ordered by means of an abstract algorithm. As VanPatten (1994) has pointed out, such grammars cannot resemble natural language grammars, because, among other things, they do not bear any referential or social meaning. The process of relating meaning to form, central in natural language learning, is therefore completely absent in the process of acquiring a finite-state grammar. Artificial languages for L2 acquisition studies, however, should bear referential and social meaning and can as such be used "for communicative purposes, . . . to regulate the social behavior of others, and . . . as instruments for other types of learning and processing of information" (VanPatten, 1994, p. 30). Similar remarks were made by Yang and Givón (1993) for their Keki language experiment and by DeKeyser (1995) for his Implexan language.

The Artificial Language eXperanto

In the present study, a computer-assisted self-study course was developed for the artificial language eXperanto. For this purpose, the existing artificial language Esperanto was chosen and modified in several respects. First of all, some of the vocabulary and spelling were changed in order to make word recognition and writing as intuitive as possible for native Dutch participants, still maintaining the feeling of dealing with a genuine foreign language. Subsequently, the Esperanto grammatical structure was modified according to the research design, for the complexity variable in morphology and syntax. The target structures resemble Spanish morphology and syntax as much as possible, in order to make the results optimally comparable with the results of a second experiment to be conducted in Spanish. This twin approach aims at combining a more easily controllable artificial language learning study with a more valid and therefore more easily generalizable natural language learning study (see Hulstijn, 1989, and Hulstijn & de Graaff, 1994, for a more detailed motivation).

Target structures 1–4 below represent the four combinations of the two variables complexity and morphology/syntax manipulated in this study (cf. Hypotheses 2 and 3). Morphology/syntax was operationalized as inflection on the one hand, represented in target structures 1 and 2, and word order phenomena on the other, represented in structures 3 and 4. Complexity was operationalized as the number

of grammatical concepts that have to be taken into account for correctly processing or producing the specific structure. To process the simple structures 1 and 3, only one grammatical concept has to be taken into account. For the noun plural (structure 1), the vowel of the stressed syllable has to be checked; the position of the negation (structure 3) depends on the word order with respect to the verb phrase. Processing the complex structures 2 and 4, on the other hand, requires the successful application of two grammatical concepts. For the imperative mode (structure 2), the correct form depends on formality as well as on affirmativity; for the object position (structure 4), the correct word order depends on nominality versus pronominality as well as on object topicalization.

1. *Plural Noun Form.* In eXperanto singular nouns end in *-o* and plural nouns in *-os* or *-es:* nouns having a back vowel (*-a-, -o-, -u-*) in the penultimate, stressed syllable add *-os;* nouns having a front vowel (*-e-, -i-*) add *-es*.[1] Examples are as follows:

(1a)	*najbaro*	*najbaros*
	"neighbour"	"neighbours"
(1b)	*turisto*	*turistes*
	"tourist"	"tourists"

2. *Inflection of the Imperative Mode.* A distinction was made between informal (2nd person) and formal (3rd person) forms, as well as between affirmative and negative forms. More specifically, the informal affirmative form ends in *-a,* the informal negative in *-us;* the formal affirmative as well as negative forms both end in *-u*. Examples are as follows:

(2a) *Ira per dia strato.*
Go (informal imperative) through that street.
"Go through that street."

(2b) *Iru per dia strato.*
Go (formal imperative) through that street.
"Go through that street."

(2c) *Nit irus trans lo plazo.*
Not go (informal imperative) over the square.
"Don't go over the square."

(2d) *Nit iru trans lo plazo.*
Not go (formal imperative) over the square.
"Don't go over the square."

3. *Position of the Negation Forms.* Any negation is formed by means of the particle *nit* in front of the inflected verb form. The negation *no/not* is expressed by *nit* only. In other negations (such as *no one, nothing, none, not at all, no more, never,* and *nowhere*), the lexical part of the negation is placed after the verb phrase, in co-occurrence with *nit* in front of the inflected verb form, thus forming a dual negation. Examples are as follows:

(3a) *Nit povak trovi mia ombrelo.*
Not can (1st person sing.) find my umbrella.
"I can't find my umbrella."

	simple	complex
morphological	***plural noun ending*** 1. front/back stem vowel	***imperative mode*** 1. formal/informal 2. affirmative/negative
syntactic	***negation position*** 1. finite+infinite verb	***object position*** 1. nominal/pronominal 2. ±stress/topicalization

Figure 1. Division of the target structures according to the variables complexity and morphology/syntax.

(3b) *Nit gin povak trovi nenier.*
Not it-OBJ can (1st person sing.) find nowhere.
"I can't find it anywhere."

4. Position of the Object. Basic word order is SVO. Unstressed pronominal objects are cliticized immediately in front of the inflected verb form. Stressed pronominal and nominal objects are placed in topic position at the beginning of the sentence, in co-occurrence with the object clitic. All animated objects are preceded by the particle *a*. Examples are as follows:

(4a) *Havas vidi a mia frato?*
Have (2nd person sing.) seen my brother-OBJ?
"Have you seen my brother?"
(4b) *Nej, nit lin havak vidi.*
No, not him-OBJ have (1st person sing.) seen.
"No, I haven't seen him."
(4c) *Dia libro gin volak aceti.*
That book it-OBJ want (1st person sing.) to buy.
"That's the book I'd like to buy."

Figure 1 summarizes the selection of the target structures for the complexity and morphology/syntax variables.

METHOD

Participants

Fifty-six participants were selected from a sample of 200 undergraduate students at the Vrije Universiteit Amsterdam who had responded to an advertisement in the university weekly calling for paid volunteers. Only monolingual native speakers of Dutch were selected, from any academic field other than linguistics. They had all had 4–6 years of education in English, German, and French, that is, a regular course

of studies in Dutch secondary education, but no experience in any other modern foreign or artificial language. From this selection, two random samples were taken of 28 participants each. Participants were paid 10 guilders per hour (approximately $6.50) for participating in the experiment.

Materials

Aptitude Test. Participants were pretested on several aspects of language aptitude and language learning capacity in order to account for within-groups variability on the posttests. Language aptitude was assessed by means of the Words in Sentences subtest of the Verbale Aanleg Test (the Dutch version of the *MLAT,* Part IV [Carroll & Sapon, 1959]; Drenth & van Wieringen, 1969) and a translated version of the Paired Associates subtest, Part V, of the *MLAT*. The Words in Sentences subtest measures sensitivity to grammatical structure, based on the identification of the analogy between the grammatical function of underlined words in two sentences. The Paired Associates subtest measures the rote memory aspect of verbal learning. Additionally, a test on language learning capacity was developed, measuring the capacity to infer the meaning of eXperanto words from context. This test consisted of 30 items, in which the correct Dutch equivalent of the highlighted word in the eXperanto stimulus sentence had to be given. A reliability analysis showed that the test was reliable to be used for further analyses (Kuder-Richardson 20 = .91 for the total test, from .83 up to .87 for the four test parts).

Instructional Packages. Both experimental groups followed a self-study course in the artificial language eXperanto. The course was provided by means of a computer program on 486SX personal computers. The program was developed using the authoring system TAIGA (1987) and had been tried out on four individuals who did not take part in the experiment. The course consisted of 10 lessons of approximately 1.5 hours each. For both experimental groups, every lesson contained the following sections:

1a. Introduction of a communicative setting by means of several short dialogs in eXperanto with meaning comprehension activities.
1b. Provision of the Dutch translation of the dialogs.
2. Vocabulary activities, consisting of translating on average 40 new lexical items from eXperanto to Dutch and from Dutch to eXperanto within sentence context.
3a. Form–meaning connection activities in which the specific function of the target structures was processed without requiring production of the structure.
3b. Production activities in which blanks had to be filled in with target structures.
3c. Sentence-level production activities based on a functional clue in Dutch.

After every item of each activity, participants received immediate feedback concerning the correctness of their response and a provision of the correct answer. At the end of each section, all items in which errors had been made were repeated once. This gave the participants a second opportunity to practice the items with

which they had difficulties. At the same time, it ensured that the participants would not rush through the lesson in order to save time, as they then would make more errors and, as a result, would have to repeat more items. During all activities, participants had access to an online bilingual lexicon. Throughout the course, the target structures were treated among other input material in situational settings, in order to increase the authenticity of the course material and to avoid too obvious a focus on the target structures. The target structures were presented in small structured doses.

The instructional packages differed for the explicit and the implicit condition in the following way: Participants in the explicit group received explanation on grammatical structures after the dialogs and comprehension activities (1a). The purpose of this explanation was to focus the participants on the target forms in the input, on their morphological or syntactic structure, and on their specific meaning. In section 1b, features of the target structure were highlighted within the dialogs. In sections 3a, 3b, and 3c, a short grammatical explanation was added to the feedback on every item. In the implicit condition, on the other hand, participants were provided a rehearsal of some example sentences in an unordered way after the dialogs and comprehension activities (1a), in order to make sure that the number of appearances of the target structures was identical for participants under both conditions. Response feedback in section 3 was not supplemented with short grammatical explanations. However, as participants in both conditions were shown the correct answer after each item, they were exposed to the same amount of input of the target structures.

Proficiency Tests. Participants were tested three times during the experiment, once halfway and two times after treatment. The proficiency tests consisted of the following four parts:

- A 60-item sentence judgment task, containing six instances of each structure two times (a grammatical and an ungrammatical version) and six filler items two times. Participants were instructed to carry out this task as quickly as possible. This direction appeared on the screen for every item.
- A 60-item gap-filling task without time pressure, containing 12 instances of each structure. Each item consisted of an eXperanto sentence with a blank and a full Dutch translation underneath.
- A 30-item contextualized Dutch-eXperanto vocabulary translation task, used for measuring participants' vocabulary knowledge and also as a distraction activity.
- A 45-item sentence judgment and correction task. Thirty items were identical to items of the first sentence judgment task; 15 were new. Participants were instructed to take all the time they needed to judge the sentences. After judgment, they were asked to correct the sentence if they had judged it as incorrect and to type it over if they had judged it as correct.

The test was administered on the computer in the order mentioned here. It contained the same items at all three sessions, except for the filler items. The items were randomly reordered within the test parts for each test session.

The test design is motivated by the extent to which the test parts allow for monitoring during processing (Hypothesis 4). The first task was carried out under time pressure, as participants were instructed to react as fast as possible, which makes explicit knowledge less accessible. The other tasks were carried out without time pressure, which makes reliance on explicit knowledge possible.

Reliability of the different test parts per test session was examined and found to be sufficient (from Kuder-Richardson 20 = 0.82 up to 0.93, except for the judgment task under time pressure at the first test session [0.59], which was due to performance on chance level for a considerable number of participants). Examples of the test items are given in the Appendix.

In addition, at the end of the experiment, oral interviews were held with all participants individually. These interviews focused on the participants' capacity to explicitly formulate the target language structures. Furthermore, they were used in order to track participants' appreciation of the learning procedure.

Procedures

For practical reasons, both experimental groups were split up in two subgroups of 14 participants each. All sessions were held in the university computer laboratory. Participants were unaware of the purpose of the experiment, nor did they know what differences there were between the instructional packages of the experimental groups until they were debriefed after the experiment. During the first session, the aptitude test was administered using a paper-and-pencil procedure. In the following weeks, the participants studied lessons 1–10, at a rate of two lessons per week. The proficiency tests were administered in the sessions following the 5th and the 10th lesson (T1 and T2, respectively). Five weeks after the second proficiency test, the test was administered again, as a delayed posttest (T3).

Two participants dropped out during the course of the experiment, one from the explicit group and one from the implicit group, leaving 27 participants in each experimental group. One participant could not attend the delayed posttest; his scores on this test were coded as missing values.

Analyses

In the analyses the following independent variables were distinguished:

Instruction, a two-level between-groups variable, distinguishing between the explicit and the implicit modes of instruction (cf. Hypothesis 1).

Complexity, a two-level within-groups variable, distinguishing between the two simple and the two complex target structures (cf. Hypothesis 2).

Morphology/syntax, a two-level within-groups variable, distinguishing between the two morphological and the two syntactic target structures (cf. Hypothesis 3).

Time pressure, a two-level within-groups variable, distinguishing between the sentence judgment tasks with and without time pressure (cf. Hypothesis 4).

Test session, a three-level within-groups variable, distinguishing between the midtest, the immediate posttest, and the delayed posttest.

Task type, a four-level within-groups variable, distinguishing between the grammar judgment task with time pressure, the grammar judgment task without time pressure, the gap-filling task, and the correction task.

Dependent variables were the sets of scores on all tests and test parts. Average aptitude scores served as an additional dependent variable in the analyses for Hypothesis 1.

First, mean scores per target structure, test session, and task type were calculated. Descriptive statistics were computed in order to check the data for normal distribution. Concerning the relationship between the effect of instruction and participants' individual language aptitude, Pearson's correlations were calculated between the average aptitude scores on the one hand and the proficiency test scores for each task type and each test session on the other, for the implicit and the explicit groups separately. To test Hypothesis 1, concerning the effect of instruction on test performance, an analysis of variance with repeated measures (ANOVA) was performed comparing the explicit and implicit groups on the individual mean scores across task types, test sessions, and target structures. Subsequently, a model was fitted in order to test whether the correlation between aptitude scores and proficiency test scores was equal for the explicit and the implicit group.

To test Hypotheses 2 and 3, concerning the differential effect of instruction depending on complexity and morphology/syntax, respectively, the individual mean scores per target structure, test session, and task type were subjected to an ANOVA with repeated measures. To test Hypothesis 4, the time pressure effect was assessed by means of an ANOVA with repeated measures comparing the mean scores per target structure and test session for the grammaticality judgment tasks with and without time pressure.

Analyses were carried out in SPSS-PC, except for those for the second part of Hypothesis 1, which were carried out in LISREL (Jöreskog & Sörbom, 1986). The alpha level was set at .05, unidirectional, for all ANOVAs.

RESULTS[2]

Table 1 shows the mean scores and standard deviations per target structure for each test session and task type. The scores were checked on normality of distribution and found to be normally distributed.

Correlations were computed between the proficiency scores in the three test sessions and the language aptitude subtest scores, for each group separately. Correlations tended to be significant for the Words in Sentences and the Word Meaning Inference subtests and nonsignificant for the Paired Associates subtest. Although the Paired Associates scores contributed little to the total correlation with the proficiency test scores, they did not correlate in a negative way. For this reason, it was decided to collapse all three subtests into a single mean aptitude score. Correlations with this mean aptitude score are given in Table 2. Correlations tend to be significant

Table 1. Means and standard deviations per target structure for the midtest (T1), immediate posttest (T2), and delayed posttest (T3), under explicit and implicit conditions

		T1				T2				T3			
		Explicit		Implicit		Explicit		Implicit		Explicit		Implicit	
Task Type	Target Structure	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)
1	Simp/morph	.62	(.17)	.51	(.08)	.70	(.19)	.56	(.16)	.75	(.20)	.62	(.16)
	Comp/morph	.66	(.17)	.50	(.14)	.78	(.19)	.69	(.13)	.71	(.19)	.62	(.18)
	Simp/synt	.67	(.18)	.57	(.17)	.73	(.21)	.71	(.21)	.74	(.23)	.75	(.21)
	Comp/synt	.50	(.14)	.46	(.09)	.75	(.13)	.64	(.14)	.77	(.16)	.65	(.20)
2	Simp/morph	.70	(.19)	.53	(.13)	.82	(.15)	.63	(.18)	.80	(.18)	.59	(.20)
	Comp/morph	.69	(.24)	.58	(.21)	.74	(.25)	.65	(.21)	.69	(.21)	.62	(.20)
	Simp/synt	.65	(.25)	.64	(.22)	.74	(.20)	.73	(.19)	.75	(.27)	.80	(.20)
	Comp/synt	.52	(.13)	.45	(.13)	.80	(.18)	.68	(.18)	.76	(.16)	.69	(.16)
3	Simp/morph	.69	(.27)	.50	(.20)	.84	(.18)	.65	(.16)	.81	(.23)	.57	(.24)
	Comp/morph	.51	(.21)	.41	(.24)	.59	(.26)	.57	(.23)	.50	(.29)	.51	(.22)
	Simp/synt	.49	(.34)	.51	(.31)	.76	(.26)	.74	(.23)	.69	(.31)	.72	(.25)
	Comp/synt	.23	(.13)	.20	(.11)	.62	(.25)	.52	(.22)	.55	(.25)	.46	(.25)
4	Simp/morph	.64	(.21)	.40	(.16)	.78	(.22)	.54	(.21)	.76	(.26)	.53	(.25)
	Comp/morph	.62	(.29)	.38	(.24)	.64	(.30)	.53	(.27)	.58	(.27)	.48	(.25)
	Simp/synt	.57	(.32)	.42	(.28)	.69	(.26)	.59	(.26)	.65	(.32)	.69	(.31)
	Comp/synt	.32	(.13)	.27	(.14)	.77	(.18)	.61	(.23)	.71	(.22)	.60	(.22)

Note: $n = 27$ for both groups. Simp/morph = simple morphological; Comp/morph = complex morphological; Simp/synt = simple syntactic; Comp/synt = complex syntactic. Task type 1 = judgment task with time pressure; task type 2 = judgment task without time pressure; task type 3 = gap-filling task; task type 4 = correction task.

at the .01 level at the immediate and delayed posttest, at all task types except for the judgment task without time pressure, for both the explicit and the implicit group.

Hypothesis 1

To check whether the explicit and the implicit group could be considered similar in terms of language aptitude, a one-way ANOVA was conducted on the language aptitude mean scores. This analysis yielded no significant results, $F = 0.41$, $df = 1$, $p = .52$. Thus, assignment of participants to the experimental groups had resulted in group similarity in terms of language aptitude.

To test Hypothesis 1, an Instruction × Test Session × Task Type ANOVA with repeated measures was conducted. The analysis showed significant main effects for instruction, $F = 8.31$, $df = 1$; test session, $F = 151.35$, $df = 2$; and task type, $F = 31.33$, $df = 3$. No significant interaction was found between instruction and test session nor between instruction and task type. Results are illustrated for the judgment tasks with and without time pressure in Figure 2 and for the gap-filling task and the correction task in Figure 3.

In order to address the hypothesis that language aptitude does not differentially

Table 2. Correlations (Pearson's *r*) between the language aptitude mean scores and the mean scores per task type and test session, under explicit and implicit conditions

Task Type	Test Session	Explicit ($n = 27$)	Implicit ($n = 27$)
1	T1	.38	.15
	T2	.47*	.42*
	T3	.50*	.55*
2	T1	.21	.02
	T2	.34	.34
	T3	.39	.34
3	T1	.52*	.32
	T2	.56*	.50*
	T3	.54*	.39
4	T1	.19	.36
	T2	.45*	.51*
	T3	.40	.50*

Note: Task type 1 = judgment task with time pressure; task type 2 = judgment task without time pressure; task type 3 = gap-filling task; task type 4 = correction task. T1 = midtest; T2 = immediate posttest; T3 = delayed posttest.
*$p < .01$.

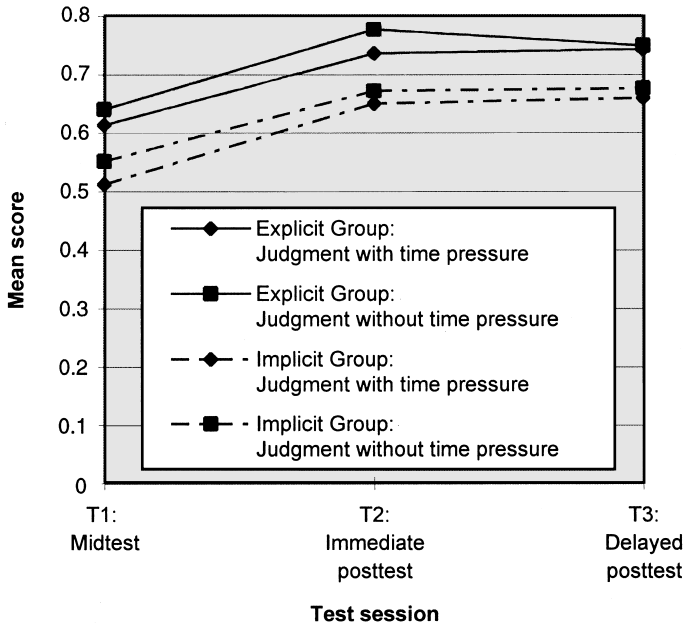

Figure 2. Proficiency test mean scores over all target structures for the judgment tasks with and without time pressure, under explicit versus implicit conditions.

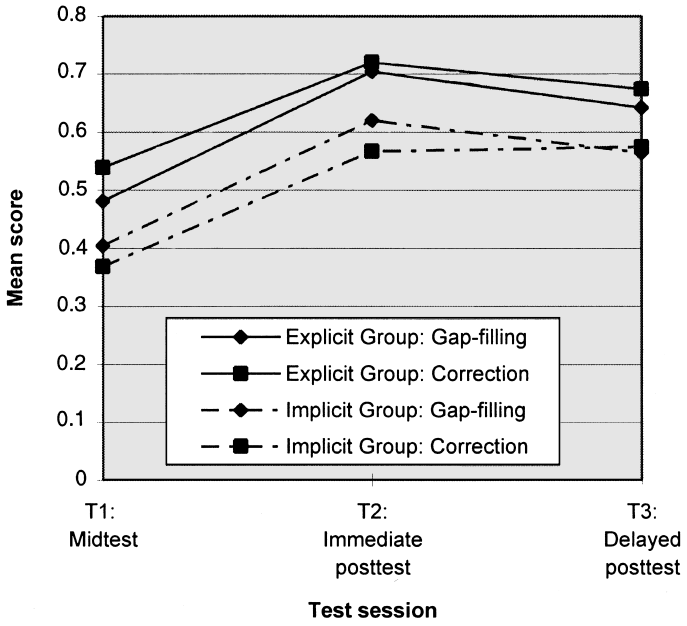

Figure 3. Proficiency test mean scores over all target structures for the gap-filling task and the correction task, under explicit versus implicit conditions.

Table 3. Chi-squares per task type comparing correlations between proficiency scores and aptitude scores under explicit and implicit conditions

	Explicit		Implicit				
	M	(*SD*)	*M*	(*SD*)	χ^2	*df*[a]	*p*
Task type 1 with aptitude	.45	(.11)	.37	(.22)	20.27	30	.909
Task type 2 with aptitude	.31	(.16)	.24	(.27)	9.18	30	1.000
Task type 3 with aptitude	.54	(.80)	.40	(.19)	18.15	30	.956
Task type 4 with aptitude	.35	(.23)	.45	(.17)	14.48	30	.992

Note: Task type 1 = judgment task with time pressure; task type 2 = judgment task without time pressure; task type 3 = gap-filling task; task type 4 = correction task.
[a]The degrees of freedom for a LISREL problem is the difference between the number of elements in the covariance matrix and the number of parameters to be estimated. Note, then, that the degrees of freedom do not reflect sample size (Marsh & Hocevar, 1985).

affect test performance under explicit versus implicit conditions, a model was fitted specifying, for each task type, equal correlations for the explicit and the implicit group between aptitude scores and individual mean scores across target structures and test sessions. For this analysis, Box' test was carried out in LISREL, which provides an overall chi-square goodness-of-fit test comparing observed and predicted correlations[3] (see Bentler & Bonett, 1980, for a discussion of this analysis).

Correlation differences and corresponding chi-squares are given in Table 3. As can be concluded from this table, chi-squares are small and *p*-values are high for all task types, which implies that the hypothesis of equal correlations for both groups cannot be rejected.

Evidence for Hypothesis 1. The hypothesis that participants receiving explicit instruction perform better on tests measuring proficiency in the target structures than those not receiving explicit instruction is supported by the data. For all three test sessions and all four task types, participants in the explicit condition scored significantly higher than participants in the implicit condition. As no interaction between instruction and test session was found, the groups appeared not to differ in growth or loss between test sessions. Furthermore, in accordance with the second part of Hypothesis 1, language aptitude was found not to affect test performance differentially under explicit as opposed to implicit conditions.

Hypotheses 2 and 3

To test the differential effect of instruction on the acquisition of simple versus complex and morphological versus syntactic structures, an Instruction × Test Session × Task Type × Complexity × Morphology/Syntax ANOVA with repeated measures was carried out. A significant main effect was found for complexity, $F = 47.26$, $df = 1$. Post hoc analyses showed that scores on simple items were significantly higher than on complex items at all test sessions. No interaction effect was found between complexity and instruction, $F = 0.16$, $df = 1$.

For morphology/syntax no main effect was found, $F = 0.13$, $df = 1$; however, a significant interaction effect existed between morphology/syntax and instruction, $F = 10.84$, $df = 1$. Post hoc analyses revealed this interaction was significant only because the scores for participants in the explicit condition at T1 were higher on the morphological structures than on the syntactic structures, whereas the scores for participants in the implicit condition at T2 and T3 were lower on the morphological structures than on the syntactic ones.

To be able to examine the effect of instruction for each of the four target structures individually, the interaction among instruction, morphology/syntax, and complexity was examined and found to be significant, $F = 6.35$, $df = 1$. From the post hoc analysis, it could be concluded that this effect was due to the scoring patterns at T2 and T3: On the simple morphological and the complex syntactic structure, participants in the implicit condition scored significantly lower than participants in the explicit condition.

Figures 4–7 illustrate the mean scores per task type on the immediate posttest for each target structure, and Table 4 reports the ANOVA results.

Evidence for Hypothesis 2. Hypothesis 2, concerning the differential effect of explicit instruction on simple as opposed to complex structures, was not confirmed. Although the complexity effect was found to be significant, no significant interaction

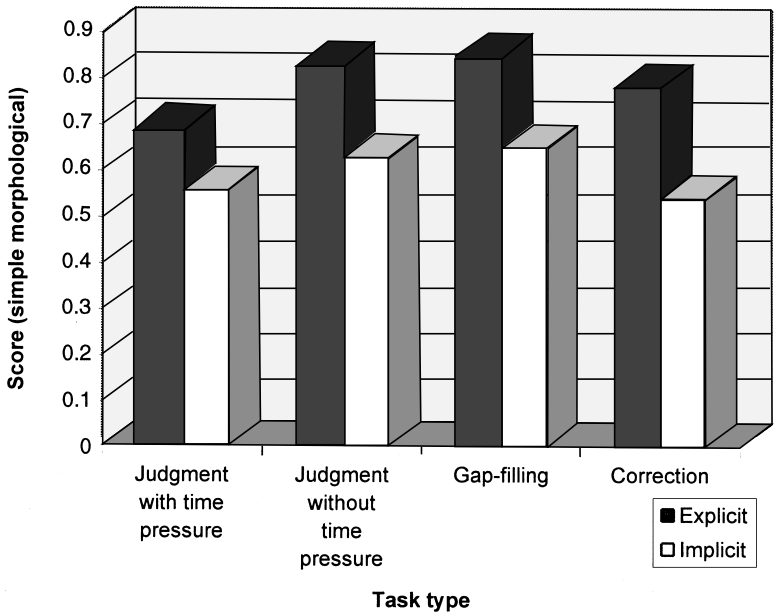

Figure 4. Proficiency test mean scores on the immediate posttest (T2) for the simple morphological structure by task type, under explicit versus implicit conditions.

between complexity and instruction could be demonstrated. This was mainly due to the participants' performance on the simple morphological structure: The difference between the implicit and the explicit group's performance was much greater than had been expected for this structure. Performance on the two syntactic structures, however, did correspond with the hypothesis: The difference between the explicit group and the implicit group, in favor of the explicit group, was greater on the complex than on the simple syntactic structure.

Evidence for Hypothesis 3. Hypothesis 3, concerning the differential effect of instruction on morphological as opposed to syntactic structures, could not be confirmed. Interestingly, however, mixed interaction effects were found: At the midtest (T1), for the explicit group test performance on the morphological structures was higher than on the syntactic structures. The implicit group scored relatively low on both morphological and syntactic structures at the midtest, which is why no significant differences were found for this group. At the immediate and the delayed posttests, however, test performance of the implicit group on the syntactic structures was found to be higher than on the morphological structures. This effect can be explained by the relatively high scores on the simple syntactic structure and the relatively low scores on the simple morphological structure for the implicit group.

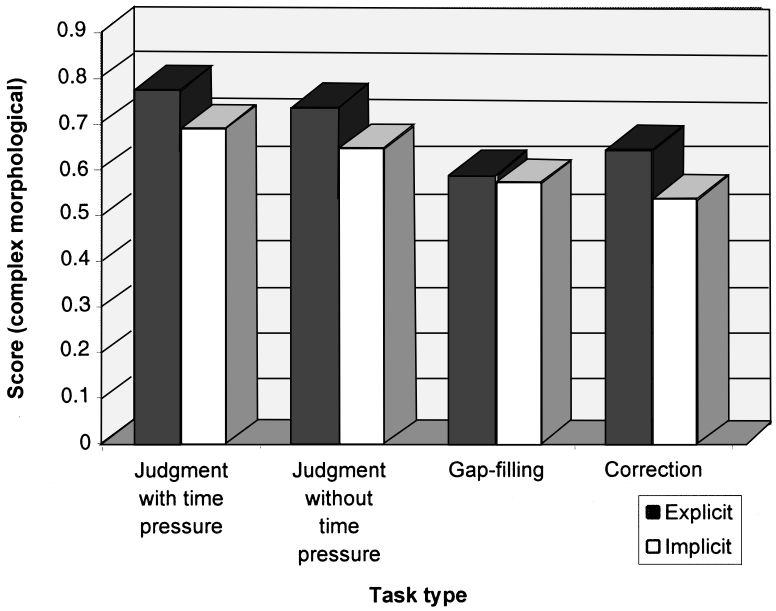

Figure 5. Proficiency test mean scores on the immediate posttest (T2) for the complex morphological structure by task type, under explicit versus implicit conditions.

Hypothesis 4

Before the differential effect of monitoring on the effect of explicit instruction could be examined, reaction times had to be compared between the grammaticality judgment tasks with and without time pressure. The mean reaction times (measured in seconds) under time pressure (task type 1) per test session were 10.0 at T1, 7.0 at T2, and 6.4 seconds at T3 for the explicit group and 9.0, 7.0, and 7.3 seconds, respectively, for the implicit group; reaction times without time pressure (task type 2) per test session were 16.2, 12.7, and 8.9 seconds for the explicit group and 14.3, 10.3, and 8.6 seconds for the implicit group. An Instruction × Time Pressure × Test Session × Target Structure ANOVA showed significant main effects for time pressure, $F = 102.9$, $df = 1$, and test session, $F = 98.1$, $df = 2$, but no significant effects for instruction, $F = 1.12$, $df = 1$, and target structure, $F = 1.15$, $df = 3$. In other words, for all target structures participants performed the judgment task under time pressure faster than the task without time pressure. This effect was equal for both groups. As this requirement was met, a comparison between groups was made for the average scores per target structure per test session of the grammaticality judgment task with versus without time pressure, by means of an ANOVA for repeated measures. Significant main effects were found for instruction, $F = 10.36$, $df = 1$, and time pressure, $F = 7.73$, $df = 1$. No significant interaction effect was found between instruction and time pressure, $F = 0.00$, $df = 1$. However, the interaction among instruction, time pressure, and target

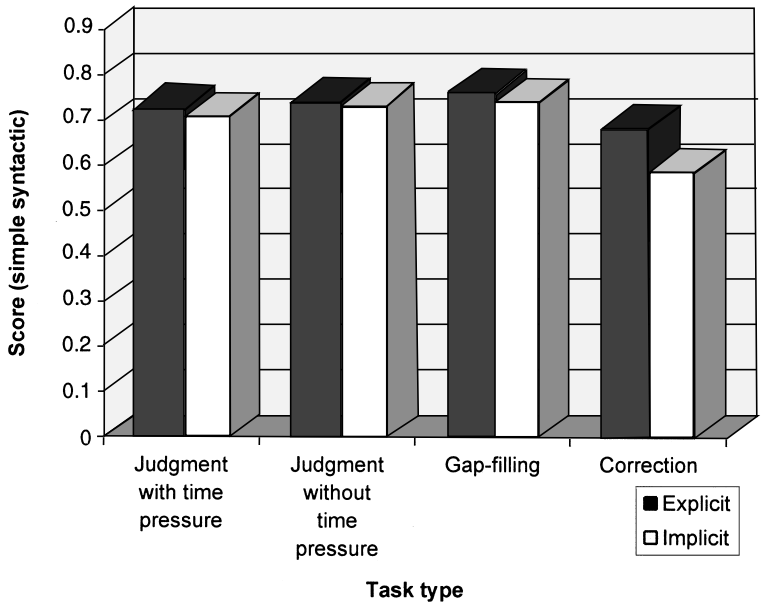

Figure 6. Proficiency test mean scores on the immediate posttest (T2) for the simple syntactic structure by task type, under explicit versus implicit conditions.

structure showed significant results, $F = 2.47$, $df = 3$. Post hoc analyses revealed this was due to the scores on the simple morphological structure. In this case, the difference between groups was significantly greater on the judgment task without time pressure than on the judgment task with time pressure.

Evidence for Hypothesis 4. The hypothesis concerning the differential effect of monitoring as facilitated or inhibited by time pressure could only partially be confirmed. Although participants performed significantly better on grammatical judgment tasks without time pressure than with time pressure, and participants in the explicit group performed significantly better than participants in the implicit group on the judgment tasks as a whole, a greater difference between groups on the task without time pressure than on the task with time pressure could only be demonstrated for the simple morphological structure.

DISCUSSION

Of the four hypotheses tested in this study, only the first could be supported unequivocally. Substantial evidence was found for the prediction that participants receiving explicit instruction during treatment perform better on tests measuring

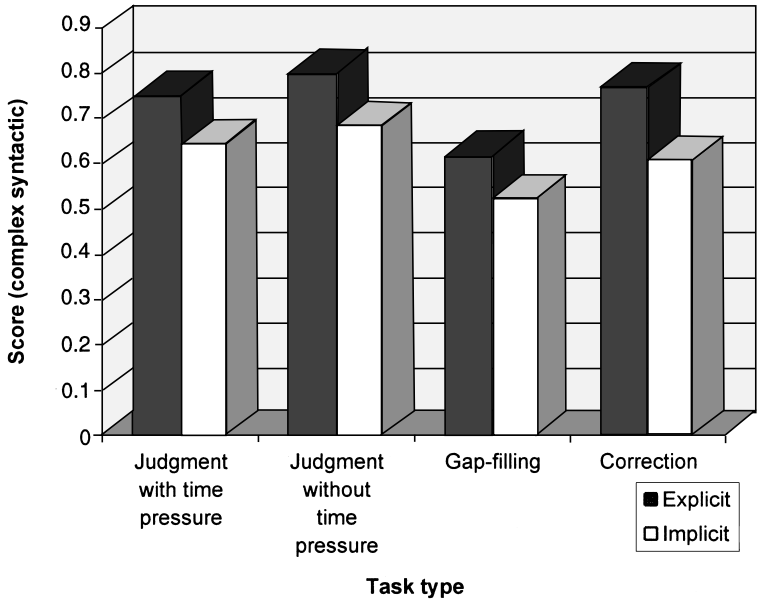

Figure 7. Proficiency test mean scores on the immediate posttest (T2) for the complex syntactic structure by task type, under explicit versus implicit conditions.

Table 4. Results of the Instruction × Test Session × Task Type × Complexity × Morphology/Syntax ANOVA with repeated measures

Source of Variation	Sum of Squares	*df*	Mean Square	*F*
Instruction (I)	7.13	1	7.13	8.31*
Test session (TS)	13.88	2	6.94	151.35*
Task type (TT)	5.90	3	1.97	31.33*
Complexity (C)	4.15	1	4.15	47.26*
Morphology/syntax (M/S)	0.02	1	0.02	0.13
I × TS	0.02	2	0.01	0.27
I × TT	0.41	3	0.14	2.18
I × C	0.01	1	0.01	0.16
I × M/S	1.31	1	1.31	10.84*
I × C × M/S	0.90	1	0.90	6.35*
I × TS × C × M/S	0.51	2	0.25	7.41*
I × TT × C × M/S	0.19	3	0.06	3.05*
I × TS × TT × C × M/S	0.07	6	0.01	0.77
Error	43.80	51	0.86	

Note: $N = 53$.
*$p < .05$.

proficiency in the target structures than participants not receiving explicit instruction. Furthermore, language aptitude was found to equally affect test performance under both explicit and implicit conditions. In both groups, higher aptitude resulted in better learning, which confirms Robinson's (1995) findings concerning the correlation between aptitude and learning under instructed, rule-search, and implicit conditions. Thus, in spite of Zobl's (1992) claim that standard deviations in proficiency measures will increase when more explicit instruction is involved, in this study the difference between "good" and "poor" language learners did not increase more in the explicit than in the implicit group. It should be stated, however, that Zobl primarily compared instructed to naturalistic learners, whereas in this study two groups of instructed learners were compared.

For Hypothesis 2, concerning the differential effect of explicit instruction on the acquisition of simple as opposed to complex structures, no clear evidence could be reported. As for the two syntactic structures, explicit instruction was much more effective for the complex than for the simple structure, as predicted by Hypothesis 2. As for the two morphological structures, however, no evidence was found: The instructional effect was not greater for the complex structure than for the simple one. This finding may have been caused by the interference of a noncontrolled variable of semantic salience. If we check the four target structures on the form–meaning connections they establish, we see that the simple morphological structure concerning the difference between two phonologically constrained plural endings is the one that most exclusively depends on form. In the case of the other three structures, however, successfully processing a specific form depends on meaning as well. VanPatten (1994) hypothesized that "learners prefer processing 'more meaningful' morphology before less or 'nonmeaningful' morphology" and that "in order for learners to process form that is not meaningful, they must be able to process informational or communicative content at no or little cost to attention" (p. 32). For the present study, this implies that learners would not easily process the difference between the two plural markers of the noun, as this difference is nonmeaningful, which would explain why explicit instruction here was at least as necessary and effective as for the meaningful imperative morphology. Complexity, then, might not only be defined in terms of the number of criteria to be taken into account in order to process a specific form, as in Hypothesis 2, but the semantic salience of those criteria should be considered as well. Future research should either keep the salience variable constant or specifically study the interaction between semantic salience and complexity.

For the interaction between instruction and morphology/syntax (Hypothesis 3), no evidence was found in this study. It could be argued that a longer or more intensive treatment would have been necessary for sufficient item-based learning to take place and for the expected interaction to reveal itself. In the treatment period of this study, target words appeared on average five times in the specific form that was tested in the proficiency tests. This might be insufficient for the acquisition of an exemplar-based representation of grammar. As long as the acquisition process has not been completed, performance on both morphological and syntactic structures might then primarily depend on nonperfect rule-based learning.[4] In this respect,

N. Ellis (1993), studying rule-based and item-based L2 learning in Welsh morphophonology, found no effect for implicit item-based learning either.

The hypothesis concerning the interaction between instruction and time pressure was not borne out by this study. Participants in both experimental conditions performed better in test conditions without time pressure than with time pressure, but an interaction effect was found only for the simple morphological structure. Again, this specific interaction could be due to semantic salience differences: Under time pressure, the participants in the explicit group ignored the specific form of the plural ending more often than the specific forms of the other target structures, as the *-os/-es* difference does not have any consequences for meaning. Participants in the implicit group largely ignored the specific form of the plural ending in both task types. Without time pressure, the participants in the explicit group had ample time to use their explicit knowledge for successfully checking the plural noun endings.

It could be argued that no overall interaction effect was found between instruction and time pressure because the test condition under time pressure still allowed for performance based on the retrieval of explicit knowledge. Although it was claimed that implicit knowledge is most directly tracked when less time is available for focusing on form, it is problematic to determine how much time pressure should be exerted in controlled experimental settings in order to eliminate reliance on explicit knowledge.[5] Although also in cognitive psychology there is no firm agreement on how to measure implicit knowledge, it is important to further develop and refine well-controlled experimental settings that aim at testing participants' performance principally based on implicit linguistic knowledge.

The findings reported in this study largely correspond with those reported in closely related studies. DeKeyser (1995) distinguished between categorical rules, which always apply, and prototypical rules, which have a much lower reliability. The rules in the present study can be considered categorical in this respect. Similar to the findings presented here, DeKeyser found that in the case of categorical rules explicit teaching is superior to having students induce the rules for themselves. Robinson (1996) compared test performance of participants in instructed, rule-search, incidental, and implicit conditions. Of these groups, the instructed group resembles the explicit group in the present study, and the rule-search group probably most resembles the implicit group in the present study. Participants in Robinson's instructed group outperformed participants in all other groups on the easy rule, and they outperformed participants in the rule-search group on the hard rule. Contrary to the findings reported in the present study, Robinson could report an interaction between rule complexity and instructional condition that approached significance.

CONCLUSION

This study has tried to test hypotheses within a research program that aims at finding evidence for a differential effect of explicit instruction on L2 acquisition (Hulstijn & de Graaff, 1994). As in previous studies by DeKeyser (1995), N. Ellis (1993), and Robinson (1995, 1996), it has been shown that it is possible to investigate

the effect of explicit instruction on L2 acquisition in a controlled experimental setting. In the present study, the simplified artificial language and the computer-controlled self-study setting allowed for acquisition to take place within a short treatment period, with clear acquisition differences among experimental groups, target structures, and individual participants.

The results of this experiment do not allow us to draw firm conclusions concerning the differential effect of explicit instruction on structures differing in complexity for morphology or syntax and on test performance with or without time pressure. However, significant differences were found between the explicit and the implicit groups for the simple morphological structure and for the complex syntactic structure. The latter finding supports the complexity hypothesis for the two syntactic structures. For the former finding, which goes counter to both the complexity and the morphology/syntax hypothesis, an alternative explanation has been suggested.

Future research carried out in well-controlled experiments as well as in more naturalistic settings, and also on nonacademic participants, should further investigate hypotheses concerning the differential effect of explicit instruction on L2 acquisition. A laboratory setting as used in the present study and in a follow-up study to be conducted in Spanish provides good opportunities for testing these hypotheses in a well-structured way: It ensures that the order and amount of input material, activities, and feedback are equal for all participants within groups, and it ensures exact control of the treatment differences between groups. The storage and analysis of responses and reaction times can provide insight not only in the product but also in the process of L2 learning.

The present study was aimed at setting up a computer-controlled language learning environment that resembles genuine self-study language courses as much as possible: naturalistic dialogs, functional topics, realistic vocabulary, diversity in activities, and exercise types. Although this setup makes the study more easily generalizable to real L2 learning environments, it might have been less controllable with respect to the frequency of exposure to specific items than other studies also carried out by means of an artificial language within a computer-controlled setting (e.g., DeKeyser, this issue; N. Ellis & Schmidt, this issue; Yang & Givón, this issue). However, the more controlled the design and the more specific the learning task, the more we bear the risk of not studying L2 acquisition anymore, but only participants' capacity to carry out some kind of cognitive puzzle. Only a careful combination of the advantages of both realistic and optimally controlled L2 learning environments can provide real opportunities for studying the effect of instruction on L2 acquisition.

NOTES

1. In Spanish, plural nouns end in *-os, -as,* or *-es,* depending on the singular noun ending. In eXperanto, all singular nouns end in the same vowel in order to make vocabulary acquisition easier. For that reason, a different criterion had to be created for noun pluralization. It is acknowledged, however, that this may affect the ease of acquisition of the pluralization.

2. In this section, only the main results of this study are reported. Detailed analyses will be reported in the author's doctoral dissertation.

3. In contrast to traditional significance testing, in this type of analysis a nonsignificant chi-square is usually preferred. When the difference is small, a small chi-square and consequently a high *p*-value will be found. Such a finding indicates that the predicted model is congruent with the observed data.

4. Jordens (1990a, 1990b) suggests that some kind of syntagmatic pattern learning might take place in the acquisition of syntactic structures. Thus, the acquisition of these structures may not be essentially different in this respect from the acquisition of morphological structures. A similar claim was recently made by Bley-Vroman (1996).

5. In DeKeyser's (1995) study, for example, participants were given 5 seconds per item in a grammatical judgment task, which was claimed to leave "very little time to draw on explicit knowledge" (p. 397). The choice of this specific period of 5 seconds had been established through a pilot study (DeKeyser, personal communication).

REFERENCES

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention & awareness in foreign language learning* (pp. 259–302). Manoa, HI: Second Language Teaching & Curriculum Center.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369–406.

van Baalen, T. (1983). Giving learners rules: A study into the effect of grammatical instruction with varying degrees of explicitness. *Interlanguage Studies Bulletin Utrecht, 7,* 71–100.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Bley-Vroman, R. (1996, May–June). *Conservative pattern accumulation in foreign language learning*. Paper presented at EUROSLA 6, Nijmegen, The Netherlands.

Cadierno, T. (1995). Formal instruction in processing perspective: An investigation into the Spanish past tense. *The Modern Language Journal, 79,* 179–194.

Carr, T. H., & Curran, T. (1994). Cognitive factors in learning about structured sequences: Applications to syntax. *Studies in Second Language Acquisition, 16,* 205–230.

Carroll, J. (1965). The prediction of success in foreign language training. In R. Glaser (Ed.), *Training, research and education* (pp. 87–136). New York: Wiley and Sons.

Carroll, J. (1981). Twenty-five years of research on foreign language aptitude. In K. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Rowley, MA: Newbury House.

Carroll, J., & Sapon, S. (1959). *Modern Language Aptitude Test—Form A*. New York: The Psychological Corporation.

Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition, 15,* 357–386.

DeKeyser, R. M. (1993). The effect of error correction on L2 grammar knowledge and oral proficiency. *The Modern Language Journal, 77,* 501–514.

DeKeyser, R. M. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17,* 379–410.

DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19,* 195–222.

DeKeyser, R. M. (in press). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition*. New York: Cambridge University Press.

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition, 13,* 431–469.

Drenth, P., & van Wieringen, J. (1969). *Verbale Aanleg Test '69* [Verbal Aptitude Test]. Lisse, The Netherlands: Swets & Zeitlinger.

Elliot, A. R. (1995). Foreign language phonology: Field independence, attitude, and the success of formal instruction on Spanish pronunciation. *The Modern Language Journal, 79,* 530–542.

Ellis, N. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology, 5,* 289–318.

Ellis, N., & Schmidt, R. (1997). Morphology and long-distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition, 19,* 145–172.

Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Blackwell.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly, 27,* 91–113.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Fotos, S. (1993). Consciousness raising and noticing through focus on form: Grammar task performance vs. formal instruction. *Applied Linguistics, 14,* 385–407.

Gardner, R., & MacIntyre, P. (1992). A student's contributions to second language learning. Part 1: Cognitive variables. *Language Teaching, 25,* 211–220.

Harley, B. (1989). Functional grammar in French immersion: A classroom experiment. *Applied Linguistics, 19,* 331–359.

Herron, C., & Tomasello, M. (1988). Learning grammatical structures in a foreign language: Modelling vs. feedback. *The French Review, 61,* 910–923.

Hulstijn, J. H. (1989). Experiments with semi-artificial input in second language acquisition research. In *Scandinavian Working Papers on Bilingualism, 8: Language Learning & Learner Language* (pp. 28–40). Stockholm: University of Stockholm.

Hulstijn, J. H. (1997). Introduction. Second language acquisition research in the laboratory: Possibilities and limitations. *Studies in Second Language Acquisition, 19,* 131–144.

Hulstijn, J. H., & de Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review, 11,* 97–113.

Jordens, P. (1990a). Grammatica in taalgebruik [Grammar in language use]. *Levende Talen, 456,* 489–496.

Jordens, P. (1990b). The acquisition of verb placement in German. *Linguistics, 28,* 1407–1448.

Jöreskog, K. G., & Sörbom, D. (1986). *LISREL: Analysis of linear structural relationships by the method of maximum likelihood.* Uppsala, Sweden: University of Uppsala.

Krashen, S. D. (1981). *Second language acquisition and second language learning.* New York: Pergamon Press.

Krashen, S. D. (1982). *Principles and practice in second language acquisition.* New York: Pergamon Press.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications.* New York: Longman.

Lightbown, P., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching. *Studies in Second Language Acquisition, 12,* 429–448.

Long, M. H. (1991). A design feature in language teaching methodology. In K. de Bot, R. Grimsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: Benjamins.

Long, M., & Robinson, R. (in press). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition.* New York: Cambridge University Press.

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin, 97,* 562–582.

McLaughlin, B. (1990). Restructuring. *Applied Linguistics, 11,* 114–128.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning, 23,* 135–158.

O'Malley, J., & Chamot, A. (1990). *Learning strategies in second language acquisition.* New York: Cambridge University Press.

Pica, T. (1983). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning, 33,* 465–497.

Reber, A. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology, 118,* 219–235.

Robinson, P. (1995). Aptitude, awareness, and the fundamental similarity of implicit and explicit second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 303–358). Manoa, HI: Second Language Teaching & Curriculum Center.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition, 18,* 27–68.

Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition, 19,* 223–248.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11,* 129–158.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11,* 11–26.

Schwartz, B. D. (1993). On explicit and negative data affecting competence and linguistic behavior. *Studies in Second Language Acquisition, 15,* 147–163.

Scott, V. M. (1989). An empirical study of explicit and implicit teaching strategies in French. *The Modern Language Learning Journal, 73,* 14–22.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15,* 165–179.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84,* 127–190.

Skehan, P. (1989). *Individual differences in second-language learning.* London: Edward Arnold.

Slobin, D. (1985). Crosslinguistic evidence for the language making capacity. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 2, pp. 1157–1249). Hillsdale, NJ: Erlbaum.

TAIGA [Twente Advanced Interactive Graphic Authoring System; computer program]. (1987). Enschede, The Netherlands: Onderwijskundig Centrum Universiteit Twente.

Terrell, T. (1991). The role of grammar instruction in a communicative approach. *The Modern Language Journal, 75,* 52–63.

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16,* 183–204.

VanPatten, B. (1994). Evaluating the role of consciousness in second language acquisition: Terms, linguistic features and research methodology. *AILA Review, 11,* 27–36.

VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15,* 225–243.

VanPatten, B., & Sanz, C. (1995). From input to output: Processing instruction and communicative tasks. In F. Eckman, D. Highland, P. Lee, J. Mileham, & R. Rutkowski Weber (Eds.), *Second language acquisition theory and pedagogy* (pp. 169–186). Mahwah, NJ: Erlbaum.

Yan-Ping, Z. (1991). The effect of explicit instruction on the acquisition of English grammatical structures by Chinese learners. In C. J. James & P. Garret (Eds.), *Language awareness in the classroom* (pp. 255–277). New York: Longman.

Yang, L., & Givón, T. (1993). *Teaching the acquisition of L2 vocabulary: The Keki language experiment* (Tech. Rep. No. 93-11). Eugene: University of Oregon: Cognitive and Decision Sciences Institute.

Yang, L., & Givón, T. (1997). Benefits and drawbacks of controlled laboratory studies of second language acquisition. *Studies in Second Language Acquisition, 19,* 173–194.

Zobl, H. (1992). Sources of linguistic knowledge and uniformity of nonnative performance. *Studies in Second Language Acquisition, 14,* 387–403.

APPENDIX

EXAMPLES OF THE TEST ITEMS

1. **Grammaticality judgment**

Instruction with time pressure:	*Antwoord zo snel mogelijk!* "Answer as quick as possible!"
Instruction without time pressure:	*Neem zo veel tijd als nodig.* "Take all the time you need."
Correct items:	
Simple morphological	*Irak fari multa bela vojagos.* "I am going to make many beautiful trips."
Complex morphological	*Prena lo dua strato ce dekstra.* "Take (informal) the second street on your left."
Simple syntactic	*Nit volas trinki nenio?* "Don't you wanna drink anything?"
Complex syntactic	*Dia vojago gin volak fari.* "That's the trip I'd like to make."
Incorrect items:	
Simple morphological	**Irak fari multa bela vojages.* "I am going to make many beautiful trips."
Complex morphological	**Prenus lo dua strato ce dekstra.* "Take (informal) the second street on your left."
Simple syntactic	**Volas trinki nenio?* "Don't you wanna drink anything?"
Complex syntactic	**Dia vojago volak fari.* "That's the trip I'd like to make."

2. **Fill in the blanks**

Simple morphological	Ga je een reis naar de bergen maken? *Iras fari un vojago* _____*?* (*al lo montos*) "Are you gonna make a trip to the mountains?"
Complex morphological	Zegt u me niet wat ik moet doen. _____ *devak fari.* (*nit min diru kio*) "Don't tell me (formal) what I should do."
Simple syntactic	Ik ga nergens heen deze week. _____ *dia semano.* (*nit irak nenier*) "I'm not going anywhere this week."
Complex syntactic	Dat museum ga ik vandaag bezoeken. _____ *visiti hodiau.* (*dia museo gin irak*) "That's the museum I'm gonna visit today."